DOCUMENT RESUME

ED 144 099 CS 203 627

AUTHOR Sodowsky, Roland E.
TITLE Freshman Speech, Freshman Writing: A Linguistic Comparison.
PUB DATE 77
NOTE 11p.; Paper presented at the Annual Meeting of the Conference on College Composition and Communication (28th, Kansas City, Missouri, March 31-April 2, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *College Freshmen; Comparative Analysis; Composition Skills (Literary); *Educational Research; English Instruction; Higher Education; *Language Patterns; Language Usage; Linguistic Performance; *Sentence Structure; *Speech Skills; Vocabulary; *Written Language

ABSTRACT

This paper reports on a study in which the speech and the writing of college freshmen were compared. Spoken samples were gathered from classroom discussion; written samples were taken from pieces written on the discussion material in a later class session. Spoken and written samples from an "A" student, a "B+" student, a "B" student, and a "C" student and a written sample from an "F" student were assessed with regard to such language measures as number of syllables per word, number of words per sentence, percentage of simple sentences, number of embeddings per sentence, percentage of uncommon words, and use of different parts of speech. The paper discusses numerous questions raised by the study and notes implications for the teaching of composition. (GW)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED144099

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Roland E. Sodowsky

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

FRESHMAN SPEECH, FRESHMAN WRITING: A LINGUISTIC COMPARISON

Roland E. Sodowsky

The average harried composition teacher probably has, as I have, spent years honing his grading technique to the point where he can give a freshman essay a fair reading in less than six minutes. Knowing this, I am not optimistic that six dozen or so additional measuring devices for freshman themes will enthrall this audience. When my colleagues and I began this pilot comparison of freshman speech and writing, I quickly discovered that every composition teacher already knew the answers to the questions we were asking. The answers were (a) "My friend, I have observed that the spoken sentence structure of freshman students is without exception more complex than the written structure"; (b) "Of course the written is more complicated. They talk in monosyllabic grunts, every one of them"; or (c) "Sir, I have taught here for thirty-five years, and I can assure you there is not one iota of difference. They write exactly as they talk." Depending on one's point of view, both (a) and (b) can be correct, and perhaps it would be nice if (c) were. Given all the considerations which come into play with every spoken sentence--context, sender-receiver coding and decoding, physiological and psychological factors, the use of the suprasegmentals of pitch, stress, and juncture--the spoken sentence is unquestionably more complex. On the other hand, take away these considerations--that is, transcribe the spoken language, put it on paper--and the average written production will show more complexity. I suggest that the (c) assumption, that writing and speaking are the same, should be correct only in the sense that many students

CS 203 627

might be better off if they really believed the written sentence was a representation of their speech and not of some alien tongue in which they are expected to write every third class period.

We gathered the spoken samples in this study from classroom discussion held over the same material that the students wrote about in a later class session; the discussion was over a short story analyzed according to a simplified formalist approach. To the objection that the classroom is an "unnatural," contrived speaking situation, one may answer that it is no more so than is the writing situation. Perhaps the most important statistic--and problem for which I have no real answers--does not appear on the data sheet at all. It becomes quickly apparent that the weakest students, those hovering near or in the "D" category, become silent when faced with the foreboding triple audience of instructor, tape recorder, and twenty-five peers, no matter how much they chatter with their friends before and after class. Thus there are no spoken data for "D" and "F" students, but I give data for an "F" paper for comparison. The papers were graded according to what I assume to be "traditional" standards--standards for mechanics, development, organization, and content--before they were analyzed. I ask you now to consider the statistics with me, referring to the table entitled "Some Selected Measures of Student Writing and Speech."

In the syllables-per-word category, the students all had a higher count in their written production than in their spoken. This fact partly reflects the use in the spoken data of such one-syllable hesitation words as "well" and even "uh," a category which in the B+ student's speech constituted over ten percent of his production. The A student had the highest syllable-per-word count in both columns, and the F paper had the lowest. Both B students ranked above the C student in the written, but not in the spoken category.

Some Selected Measures of Student Writing
and Speech

	C		A		B+		B		F
	written	spoken	written	spoken	written	spoken	written	spoken	written
1. Syllables/word	1.32	1.25	1.45	1.30	1.36	1.24	1.41	1.28	1.27
2. Words/sentence	24.5	10.4	17.0	14.1	15.1	19.6	22.3	16.3	11.5
3. % simple sentences	40.0	56.3	25.9	36.7	47.9	29.6	36.4	52.0	62.8
4. Embeddings per sentence	5.25	1.94	3.98	1.26	3.17	1.09	6.91	2.16	1.29
4a. Number of words/ embedding occurrence	4.67	5.36	4.27	11.19	4.76	17.98	3.22	7.54	8.91
5. % different words	40.1	51.0	35.4	35.7	35.5	26.3	54.3	39.1	45.8
6. % borrowings	20.8	7.5	27.6	16.3	22.7	17.1	24.1	8.4	20.9
7. Number of usages/ borrowed word	1.36	1.15	1.73	1.48	1.99	1.91	1.59	1.46	1.66
8. Number of usages/ native word	2.66	2.08	4.28	3.38	3.21	4.77	1.94	3.06	2.59
9. different conjunctions	10	5	9	13			6	9	7
10. different prepositions	26	9	12	12			12	10	9
11. different pronouns	24	13	14	12			10	17	7

--Roland Sadowsky
Oklahoma State University

Generally, these figures are about as one might predict, if he assumes that a higher syllable count is an indicator, however weak, of more complexity. Incidentally, for comparative purposes one may note that the first 100 words of Moby Dick averages 1.59 syllables per word, of Huckleberry Finn 1.26. The count obviously cannot drop below one syllable per word, while any writing that approaches a two-syllable average is difficult to read.

In the words-per-sentence category, the data was not so cooperative, although the F paper had the lowest written average. The C paper averaged 24.5 words per sentence, approaching the average length of academic works; next came the B paper with 22.3, then the A paper with 17, and finally the B+ student's, 15.1, 4½ words shorter than his spoken average. The closest correlation here was between the spoken and written lengths for the A student, only three words apart. By far the widest divergence was in the C student's figures, where the average written sentence was nearly two and one-half times as long as his spoken. Again, for comparison, I note that the first 100 words of Moby Dick has 28.6 words per sentence, of Huckleberry Finn 18.2.

I have little faith in the traditional sentence type categories, being well aware that a so-called "simple" sentence may in fact be far more complicated than a traditionally defined compound-complex sentence; however, the data do show the F paper to have by far the highest proportion of simple sentences, and with the exception of the B+ student, the data show that the students write proportionally more complex and fewer simple sentences than they speak. The A student's totals point to more balance of all four types.

The embedding-per-sentence and number-of-words-per-embedding-occurrence categories may be the most important on the page. These figures reflect one

important measure of the student's linguistic sophistication, that is, his ability to embed one sentence within another through the use of adjectives, prepositional phrases, infinitives, gerunds, series, ellipses, appositives, or subordinate clauses. Having the linguistic capacity to combine three sentences such as "I have a house. The house is white. The house is on a hill" into one sentence, "I have a white house on a hill," probably enables the student to say more in fewer words, to achieve a denser texture in his writing; I say "probably" because, as politicians too frequently remind us, I do not think it necessarily follows that sophistication of language form means sophistication or profundity of content.

The embeddings-per-sentence figures show the F paper to rank woefully low in sentence complexity, and, combined with the 62.8% simple sentence figure, suggest that this student needs considerable writing laboratory help if he is to catch up with his peers. Line 4a, which shows how often a sentence embedding occurs, has the F student writing one embedded sentence every nine words, about half the frequency of the other students. In the complete breakdown of embedding types (not shown in this table), this paper had none or only one example in six of the nine embedding categories. The highest embedding ratio belonged to the B student, who averaged in her 22-word sentences almost seven embeddings, or an embedded sentence every 3.22 words, pointing to a very dense texture; the B+ student, conversely, with relatively short sentences, had a low written embedding ratio, perhaps correlating with his spoken ratio, which was by far the lowest of the examples. The most surprising data here appear for the C student, who in his long sentences averages over five embeddings each and uses every type except the appositive. Every student used more embeddings in the written examples, from two and a half to over three times as many. As in the sentence-type break-

down, the A paper shows a relative evenness in frequency in the use of all the categories except two, perhaps suggesting competent versatility. Note that on line 4a, frequency of embedding, the students with the highest grades showed the widest divergence in their spoken and written production.

In the various categories of "borrowed" and "native" words, "native" means words that were already in the language in Old English times. This arbitrary delineation has the effect of making very a borrowing while warp is native, but it generally distinguishes the "heart-food-and-shelter" vocabulary of English, the words children learn first, from the rest, and it includes in the native group practically every function word, that is, words such as of, the, when, which have primarily grammatical rather than lexical meaning. A better classification system would be one of the various "1,000 most commonly known words" compilations. We note that in every case the percentage of total borrowings is higher in written, most radically so for the C and B student. A significant statistic here is the frequency-of-usage-per-word category, where the A and B+ students show the highest rates of repetition for both foreign and native words, in both spoken and written production. The students who received the highest grades re-used words more often than those scoring lower. We will return to the implications of this and other correlative data later. One also should note that the three "low" papers show the highest percentages of "different" words. This figure indicates that, proportionately, the students with the lower scores, even the F student, had more vocabulary items.

The categories for percentage use of the parts of speech (I regret that figures for the B+ students are not available) show an increase in the use of adjectives and nouns for the written effort of each student while adverb usage drops for all and verb and pronoun usage decreases for the C and B

students. The relatively high percentage of verbs in the A paper is traceable to the writer's use of verbs in parallel series. The noun totals appear to increase in writing at the expense of pronouns, at least partly because in speech the students use pronouns without clear antecedents at will, but are much more cautious in writing.

Significant in the last three categories on the page are the totals showing the number of times which the C student uses different prepositions and pronouns in writing, double that of the other papers. These high figures, taken with this student's high percentage of different words, his low usage-per-word totals, the high embedding percentage, and his long sentence length, indicate a basic grading problem for the composition teacher, one which I can point out, but cannot solve. Most of us have encountered the student who sometime in high school or even earlier observed (a) that mechanical errors in essays result in low grades, then discovered (b) that short, simple subject-verb-complement sentences are unlikely to have errors in them, and so (c) wrote such sentences from then until some English teacher (d) penalized him a grade or two for choppy, monotonous sentences. The figures I have cited here identify the A student as a very careful writer, one who uses and re-uses a relatively small number of function words (nine conjunctions, twelve prepositions, fourteen pronouns, and the articles--total, 38) extremely well; the C student, on the other hand, takes all kinds of composing "risks": he used sixty different function words, writes longer sentences, tries out words not solidly established in his linguistic competence, and has a high embedding rate. And--I say in good periodic style at this point--he receives lower grades. But after this analysis there are still other factors at least as important, not yet considered. Here the element of <u>content</u> enters in. His essay does not <u>say</u> as much as the others.

There are semantic lapses more difficult to judge and explain than the formal problems; despite the apparent richness of vocabulary, there is a poverty of texture, almost endless repetition of idea, low level of interpretation. But these are not matters for this paper. What I am concerned with here is whether the teacher should tell this student to rein in his linguistic exuberance a bit, to write less (speaking now of linguistic form and variety) for a better grade. I am not sure, incidentally, that the content of this student's work would improve if he wrote more simply. I am equally concerned that the A student is at least partially rewarded for staying on a nice, safe linguistic plateau among old familiar words and constructions. If there is an answer, it may lie in the use of measuring devices such as those discussed here which help identify the linguistically daring student, help us understand why he may misuse a demonstrative pronoun or forget the number of the verb occasionally.

These figures raise, I hope, far more questions than they answer. I wonder, for example, if the increase in relative complexity in the written samples is traceable only to the facts that the writer, even in a timed classroom situation, simply has more time than when he speaks and so chooses a less familiar but possibly more specific word and deliberately builds a more complex sentence structure, or I wonder again if the increased complexity of written form is a compensation for the many other factors surrounding speech utterances. Another question: various percentages of hesitation words or phrases, "diluters of meaning" such as "kind of" and "sort of," filler phrases or sender comprehension check phrases such as "you know" make up varying percentages of speech utterances, up to 10½% for the B+ student in my sample; I wonder if there is some kind of compensation for these phrases when speaker becomes writer--if, in the case of this student, short sentences

and a low embedding rate are parallel compensatory characteristics? Other questions involve the sample-taking process: a logical step now might be to reverse the sequence--that is, let the students write first, then discuss. One would expect the various measures to be much closer together. One element I have not yet considered is the relative complexity of the verb phrases in the samples. Later work will include this measure. Another element is the question of spelling. Pronouncing a word such as brought is relatively far simpler than spelling it. Do some writers avoid such words, even though the words are assuredly present in their speaking competence?

Finally, there are questions about the intangibles in the classroom. I think most composition teachers would agree that the physically attractive student usually proves to be the more confident, and therefore more frequent, speaker. Does this fact of more speech practice in the classroom give such students a writing advantage? I suspect, but cannot prove that it does. There is also the circumstance of the all-important grade. Composition students expect to be graded for what they write and not for what they say; how much does this expectation affect what they say and write? Another consideration is the matter of sheer student interest in the classroom. The student who is bored by the discussion and does not participate as a sender, a talker, unless forced to, nevertheless gets some practice as a receiver, unless he can turn himself completely off; if the instructor manages to involve him, will he write a better essay? I think so, but I wonder if that is simply teacher ego at work. And one final question: if a student's writing ability is closely correlated with his speech, then how much can a freshman composition teacher hope to do? We know that language learning is primarily an infant and childhood activity and diminishes rapidly with age, going to something like one percent proportionately by

age nineteen. I can teach most students to avoid comma splices and to stop hanging participial phrases after semi-colons; I don't think these are matters of language learning. But how much can the student change his speech patterns? I do not mean to end on a pessimistic note. I honestly believe the composition teacher can have a role beneficial to the student in that last one percent of language acquisition. But the difficulty of the task is clear, and the teacher should use every available tool. The various measures I have discussed here probably are among the tools, and the sooner we find out what the correlations between the way students talk and the way they write are, the sooner we can be more effective teachers of writing.

--Roland Sodowsky
Oklahoma State University
Stillwater, OK 74074